

CapriDB - Capture, Print, Innovate: A Low-Cost Pipeline and Database for Reproducible Manipulation Research

Florian T. Pokorny*, Yasemin Bekiroglu*, Karl Pauwels, Judith Bütepage, Clara Scherer, Danica Kragic

Abstract—We present a novel approach and database which combines the inexpensive generation of 3D object models via monocular or RGB-D camera images with 3D printing and a state of the art object tracking algorithm. Unlike recent efforts towards the creation of 3D object databases for robotics, our approach does not require expensive and controlled 3D scanning setups and enables anyone with a camera to scan, print and track complex objects for manipulation research. The proposed approach results in highly detailed mesh models whose 3D printed replicas are at times difficult to distinguish from the original. A key motivation for utilizing 3D printed objects is the ability to precisely control and vary object properties such as the mass distribution and size in the 3D printing process to obtain reproducible conditions for robotic manipulation research. We present CapriDB - an extensible database resulting from this approach containing initially 40 textured and 3D printable mesh models together with tracking features to facilitate the adoption of the proposed approach.

I. INTRODUCTION

In this work, we focus on three fundamental steps that a robot needs to perform to manipulate an object: *object data acquisition*, *object representation* and *tracking*. During the *object data acquisition phase*, sensors are used to obtain a geometric model of the object. Typically this results in an *object representation* such as a mesh model or point-cloud which serves as input to grasping and manipulation planning algorithms such as [21, 4, 8]. In the manipulation phase, the object’s position and orientation needs to be *tracked* – initially, in order to execute a planned grasp, but also during the manipulation, for example to detect slippage. A key problem manipulation research is facing is the difficulty in reproducing results and the complexity of benchmarking in this domain. In the last decade, a large number of groups have proposed various benchmarking schemes, and several 3D object databases have been developed [5, 6, 7, 10]. However, these attempts have unfortunately only found partial adaptation in the research community. In our opinion, the following have in particular inhibited widespread adoption:

- Sensor dependent object models. Many works [6, 10] rely on costly or specially designed scanning setups. Our approach only requires a single hand held camera.
- Unavailability of objects involved: Databases such as [10] suffer from the lack of worldwide availability of the objects involved. While the authors of [6] propose to post object benchmark datasets to interested researchers,

our approach is to instead solve this problem by incorporating the use of 3D printed real world objects which can be printed and reproduced in a decentralized manner anywhere in the world.

- Size constraints: Mostly objects of size 5 to 50cm have been considered in databases [5, 6, 7, 10] due to sensor constraints, and many of these objects can be manipulated only by robotic hands of compatible size. 3D printing allows us to scale objects to within the 3D printer’s capabilities and monocular images can be taken of objects with a wide range of scales.
- Dependence on material properties: 3D printing/milling, allows researchers to study the impact of object material properties in isolation, allowing researchers to create objects in a large number of materials and with controlled mass distributions.
- Currently, databases do not provide a reference visual tracking system, resulting in a large source of error and discrepancy between experimental setups. We incorporate our approach with the state of the art and freely available real-time visual tracking system [16].

To address the above issues, we introduce a 3D object database for manipulation research as well as an associated efficient and low-cost workflow to capture, print and track new objects. Figure 1 outlines the steps of our approach. We utilize Autodesk’s 123D catch software to acquire object models. This only requires camera images of the object taken from a variety of arbitrary angles around the object without custom setup. We present details on a resulting object database containing mesh and texture information for 40 objects as well as reference tracking image frames. Our approach is complimentary to recent efforts such as the recently proposed YCB database [6] which, unlike our work, focuses on benchmarking protocols besides providing a set of objects scanned with a particular high-quality scanning rig. Additional key differences of our approach include the use of 3D printing rather than relying on the delivery of original objects, the integration with a particular tracking solution as well as the low-cost extensibility of our dataset which does not rely on a specific scanning setup or object scale but only on a hand-held camera. The database and associated documentation is hosted at <http://www.csc.kth.se/capridb/>. In this paper, we furthermore illustrate potential applications of our approach. In particular, we verify that the obtained object models can be 3D printed with texture and that the pose of these printed objects can be tracked successfully. We furthermore perform initial grasping experiments

*These authors contributed equally to the paper.
All authors are with the Centre for Autonomous Systems, CAS/CVAP, KTH Royal Institute of Technology, Sweden,
{yaseminb, fpokorny, kpauwels, butepage, cescha, dani}@kth.se

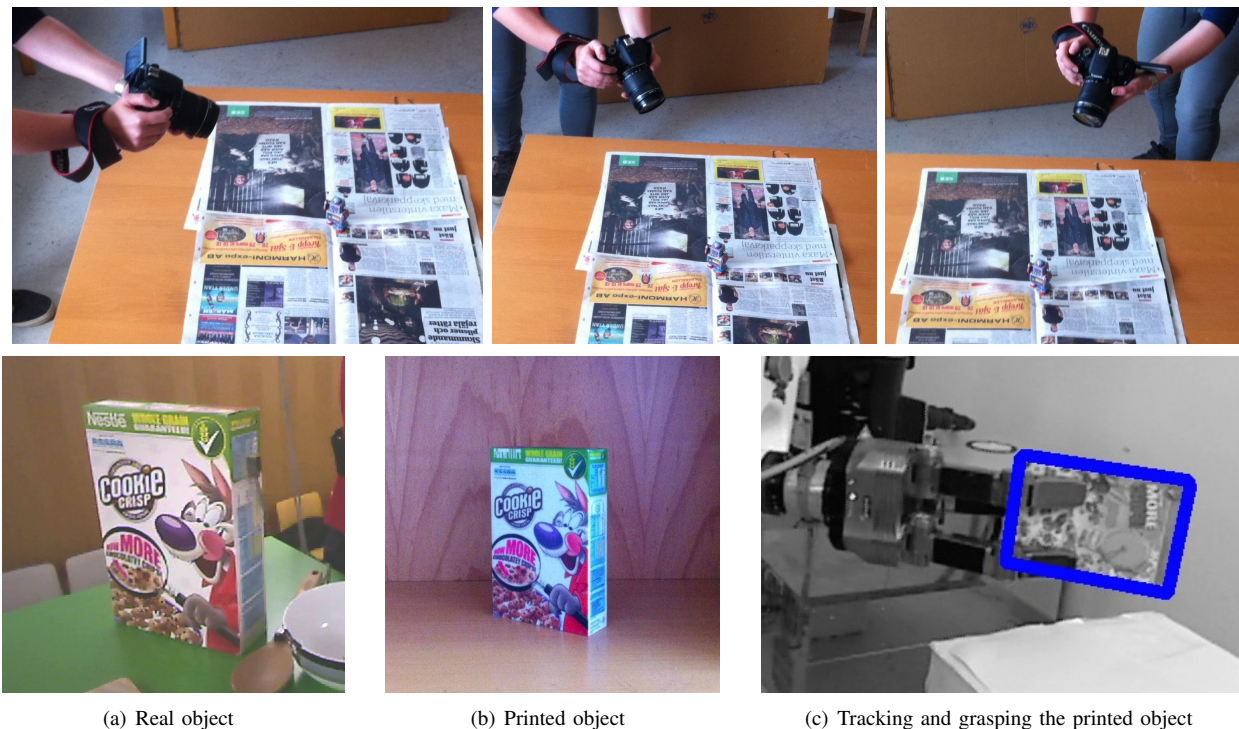


Fig. 1. Outline of the proposed data processing approach. Upper row: Images of a small toy robot (in the center of figures) are captured from various arbitrary angles with a regular camera. Lower row: An example of an acquired textured mesh-model of a cereal box is 3D printed in color. The resulting 3D printed object is tracked and grasped with a robotic Schunk SDH dexterous hand.

using estimated poses of printed objects which are calculated using the mesh-models obtained from the original real-world objects.

II. METHODOLOGY

In this section, we describe the key components of our data processing pipeline: 3D model construction, 3D printing, and tracking.

A. Textured 3D Model Construction

While current grasp databases often rely on carefully calibrated specific capturing equipments [6, 10], our approach is to use a simple digital camera in conjunction with a freely available 3D reconstruction software to capture high-quality 3D objects. This approach has recently become possible due to the availability of high-quality 3D reconstruction software relying only on monocular images. To reconstruct a 3D model from a collection of photos, we utilize the web-based free Autodesk 123D catch service [2]¹ using approximately 40 pictures of the object from various angles. To improve the quality of reconstruction, we place the objects on a textured background consisting of a collection of newspapers. Figure 2 displays a partial screenshot of the software, illustrating the automatically reconstructed camera positions. The scanned object is visible in the center of this visualization.

¹Other solutions with similar quality are available, e.g. Agisoft PhotoScan [1].

B. 3D Model Postprocessing

The acquired 3D mesh model requires post-processing in order to result in a clean and fully specified model².

Firstly, the metric dimensions of the model have to be specified in centimeters with the help of reference points for which we use the Autodesk 123D catch software. As the initially obtained 3D mesh model contains not only the object but also some parts of the surrounding environment, such as the surface on which the object might rest, these extraneous parts of the extracted mesh need to be removed. We use the open source software Meshlab [13] for this purpose. Figure 2 illustrates post-processing steps where areas that do not belong to an object are manually removed from the initial model. In the final manual processing step, holes in the mesh are closed. Holes arise, for example on the underside of the object, when the object rests on a planar surface when the photos are taken. For the hole filling, we used the open source 3D modelling software Blender [3], which also can be used for rotating and scaling the models as desired. Furthermore, we use a specific object pose tracker, which we describe later, to demonstrate that the pose of these models can be determined. The tracker requires the dimensions of the mesh model to be provided in meters, in accordance with the ROS convention. Therefore, as a final postprocessing step, the models are scaled accordingly. After this processing step, we obtain a mesh model whose geometry is stored in Wavefront OBJ format, a mesh to texture mapping stored in

²Detailed instructions regarding this process are available on the website.

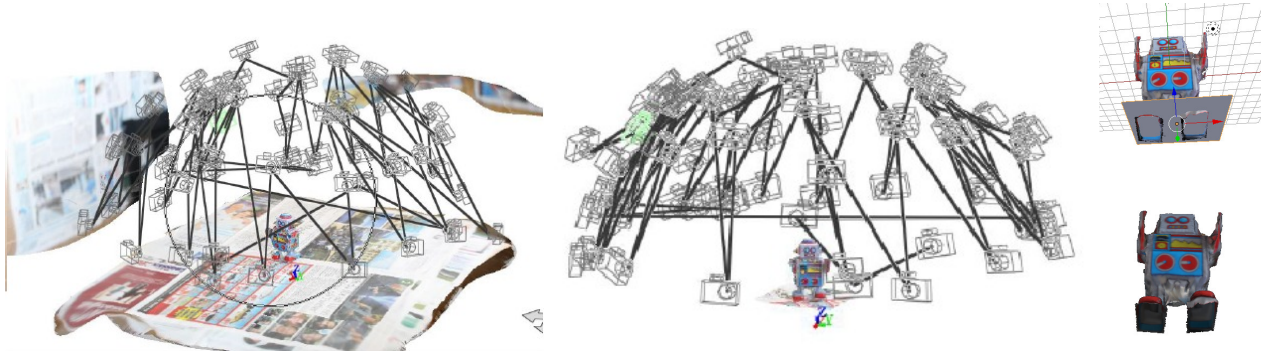


Fig. 2. Left and middle figures: Construction of the 3D model with Autodesk’s free 123D catch application. The reconstructed camera poses and the central object pose is displayed. Rightmost figures: Post-processing of the acquired textured mesh model, where the mesh is made watertight and surface areas not belonging to the object are removed.

MDL format as well as a texture file, which is stored as a JPEG image.

C. 3D Printing Textured Objects

Our goal is to make objects accessible to everyone both as 3D mesh models and in physical/graspable forms. The rapidly advancing field of 3D printing makes it possible to 3D print objects rather than obtaining originals. A large range of on-line services offer to print highly textured objects in color. This allows anyone to reproduce objects based on the provided 3D mesh models and to use these for robotic manipulation research³. We have printed several objects (see Figure 8 and Figure 4) through the company iMaterialise [9], see Section II-E.1. Note that 3D printing also enables us to scale objects as desired, vary the internal mass distribution and select a wide range of object materials. We believe this opens up promising new possibilities to study frictional and dynamic behavior in robotic manipulation in a controlled fashion and independently of shape. Figure 3 displays examples of printed objects which we scanned and printed.

D. Real-Time Tracking and Pose Estimation

We use a state-of-the-art image-based object pose estimation method that uses sparse keypoints to detect, and dense motion and depth information to track the full six degrees-of-freedom pose in real-time [16, 15]. This method achieves high accuracy and robustness by exploiting the rich appearance and shape information provided by the models in our database. This pose estimation method is publicly available as a ROS module⁴. We validate our methodology by successfully detecting and tracking the pose of printed models on the basis of the mesh models generated from the original objects. An example tracking result is shown in Figure 3 for a scene with occlusions and multiple 3D printed objects and in Figure 7, where a PR2 robot’s onboard arm camera is used to track several 3D printed objects. Both RGB and RGB-D cameras can be used with this approach.

³Some of the available printing services are Shapeways (US) [23], Cubify Cloud Print (US) [19], Sculpteo (France) [22], iMaterialise (Belgium) [9]

⁴www.karlpauwels.com/simtrack

E. Application Scenarios

Here, we highlight various potential directions that the proposed approach and database could be used for:

1) *Integrated tracking and grasp planning*: Most grasp planners [12, 14, 20, 11] rely on estimated object pose to parametrize grasps, i.e., to calculate wrist position and orientation. The models obtained with the approach proposed here (see also Figure 8) can be utilized for pose estimation as can be seen in Figure 3 and Figure 7. We have conducted experiments using the 3D printed objects displayed in Figure 8, which are based on a real PR2 robot, a cereal box, a toy robot, a toy horse and toy duck. Note that these objects have complex shapes and vary in scale. The printed objects are then tracked based on the models obtained from the original real-world objects using the tracker described in the previous section.

We have conducted preliminary experiments with some of the object models in order to demonstrate the feasibility of using these models for combined pose tracking and grasping purposes. For our grasping experiments, we used a robot composed of an industrial KUKA arm and a Schunk Dexterous (SDH) Hand with a predefined hand preshape, as displayed in Figure 5. Based on the estimated object pose, we executed side and top grasps by placing the wrist to a predefined distance from the object’s center along its vertical and horizontal axis and closing the fingers. Figure 7 displays tracking results based on a PR2 robot’s onboard arm-camera, where the robot detects and tracks the horse, cereal box and robot toy in the same scene while the cereal box is being lifted by the robot. Note that the proposed tracking system can reliably handle the resulting occlusions in this scene.

2) *Replicable Manipulation Research*: Since 3D printing has become widely available, the proposed approach enables researchers to create replicable robotic experiments by running robotic experiments using 3D printed objects. Furthermore, 3D printed objects may also serve as a controllable testing environment for tracking algorithms other than the proposed reference tracking system.

3) *Manipulation of Objects with Controlled Physical Properties*: The proposed 3D printing (or milling) process allows for various choices of materials such as plastics,

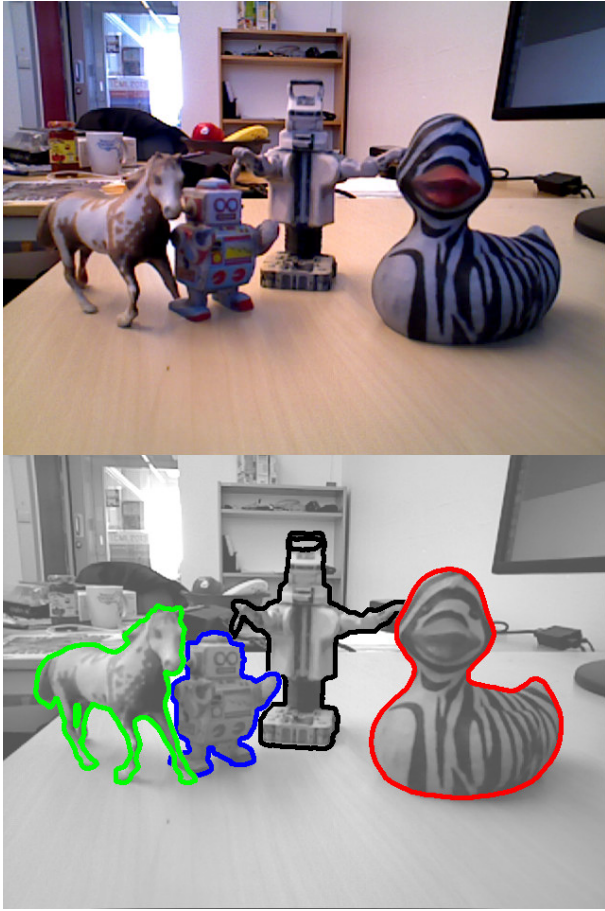


Fig. 3. Top: Examples of 3D printed objects whose 3D textured model was acquired using the proposed methodology: A horse model, a toy robot, a rescaled PR2 robot and a toy duck. Bottom: Pose tracking results of the printed objects based on textured models acquired on the originals.

sandstone and wood. Using this approach, object properties can be separated from object geometry, as specified by the scanned meshes. Furthermore, the internal mass distribution of an object can be modified by partially filling the printed meshes, or keeping them hollow, etc. This, we believe, provides an interesting avenue for robotic manipulation research to focus on sub-problems such as robustness to variations in friction coefficients, mass distribution, etc. Another important aspect is the scaling of the resulting objects. Many current robotic hands have dimensions differing from a human adult or child’s hand. By printing objects at a range of scales, researchers may be able to study the success of robotic grasps depending on scale and could, for example, optimize robot hand design with respect to object size. A further interesting direction of research is the study of grasps on continuous families of perturbations of objects. A simple example would be grasping cones with various angles at the apex to understand frictional properties, but more generally, shape and grasp *moduli spaces* ([18, 17]) defined by deformations of shapes and grasps could be studied by 3D printing perturbations of existing objects. This constitutes a direction of research we would like to investigate in future,



Fig. 4. Side by side comparison of original models (right) and 3D printed objects (left).

in particular.

III. CONTENT AND FORMAT OF INITIAL CORE DATABASE RELEASE

The initial release of CapriDB contains 40 watertight textured mesh models of the objects listed in Table III and depicted in Figure 6. Mesh models are stored in Wavefront OBJ format, a mesh to texture mapping is provided in MDL format and an associated texture file is stored as a JPEG image for each object. The objects for the IEEE ICRA Amazon Picking Challenge 2015 are also included in the database. Table III lists the physical dimensions of these objects, their weight and original material as well as additional notes which will also be stored in CapriDB. In addition, the initial database release contains the original photos (approx. 40 per object) used to construct the mesh approximation in JPEG format.

To facilitate performance evaluation, we also include reference images (in JPEG) and associated tracking boundaries (overlayed JPEG based on object poses acquired from the

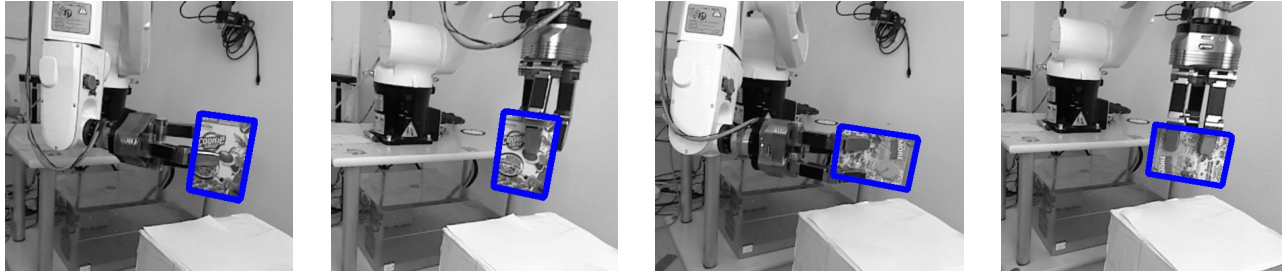


Fig. 5. Grasping experiments with the Kuka arm, Schunk Hand and the printed box: (a) Side grasp with the box when it is standing upright originally, (b) top grasp when the box is standing upright (c) side grasp when the object is lying sideways and the back side of the object is visible to the tracker, (d) top grasp when the object is lying sideways. These experiments were performed based on the object poses estimated by the tracker which used the 3D models obtained from the real objects, illustrating that the texture of the printed object matched the original texture sufficiently well. For tracking, images from a Kinect sensor were used. The object can continuously be tracked during grasping and lifting. The blue frames around the objects indicate the tracked poses.



Fig. 6. Initial set of 40 objects in the core database.

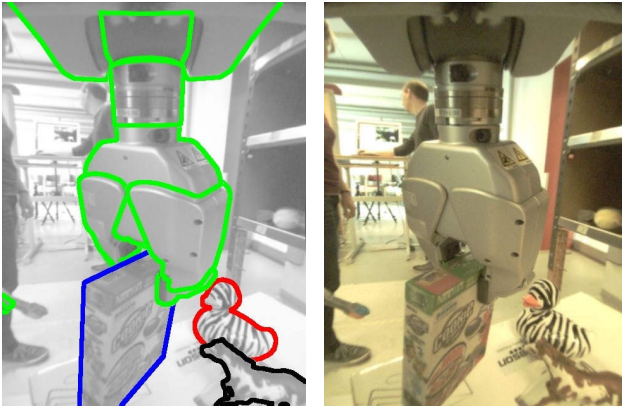


Fig. 7. Example tracking of printed objects based on the PR2 robot's arm camera.

tracker) for each object as in Figure 3 to test and compare other tracking methodologies. Figure 9 shows how the database and interactive tracking could be used for benchmarking using a pre-defined scene layouts. The included scenes and object poses can be used as ground truth to set up a system using these object models and the tracker. More information about the tracker's accuracy can be found in the work of Pauwels, Rubio, and Ros [15].

IV. CONCLUSION AND FUTURE WORK

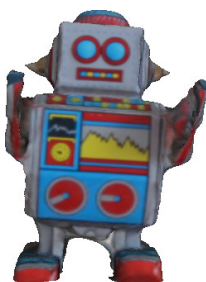
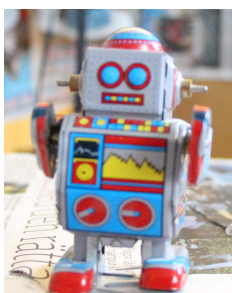
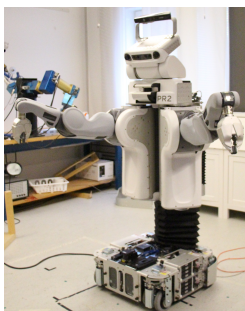
We have introduced an inexpensive pipeline which utilizes 3D printing, inexpensive mesh reconstruction from monocular images and a state of the art tracking algorithm to facilitate reproducible robotic manipulation research. Our approach only requires a regular RGB or RGB-D camera and images taken from a set of angles and access to a 3D printing service. The initial database of 40 scanned objects is available at <http://www.csc.kth.se/capridb> and we plan to continue contributing to this database over time by adding more objects and object features.

ACKNOWLEDGMENTS

This work was supported by the EU through the project RoboHow.Cog (FP7-ICT-288533) and the Swedish Research Council.

REFERENCES

- [1] AgisoftPhotoScan. <http://www.agisoft.com>. 2015.
- [2] Autodesk. <http://www.123dapp.com>. 2015.
- [3] Blender. <http://www.blender.org>. 2015.
- [4] C. Borst, M. Fischer, and G. Hirzinger. "Grasping the Dice by Dicing the Grasp". In: *IROS*. 2003, pp. 3692–3697.
- [5] I. M. Bullock, T. Feix, and A. M. Dollar. "The Yale human grasping dataset: Grasp, object, and task data in household and machine shop environments". In: *IJRR* (2014), p. 0278364914555720.
- [6] B. Calli et al. "Benchmarking in Manipulation Research: The YCB Object and Model Set and Benchmarking Protocols". In: *arXiv preprint arXiv:1502.03143* (2015).
- [7] C. Goldfeder et al. "The columbia grasp database". In: *IEEE ICRA*. IEEE. 2009, pp. 1710–1716.
- [8] K. Huebner, S. Ruthotto, and D. Kragic. "Minimum volume bounding box decomposition for shape approximation in robot grasping". In: *IEEE ICRA*. 2008, pp. 1628–1633.
- [9] iMaterialise. <http://i.materialise.com>. 2015.



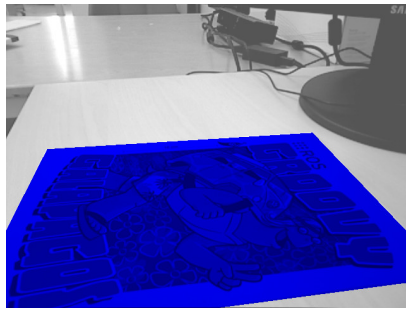
(a)

(b)

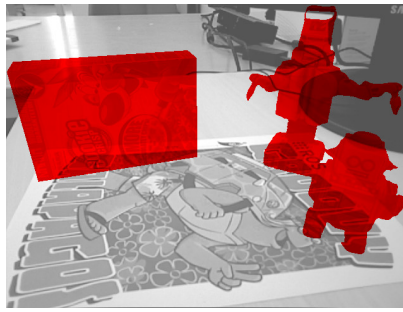
(c)

(d)

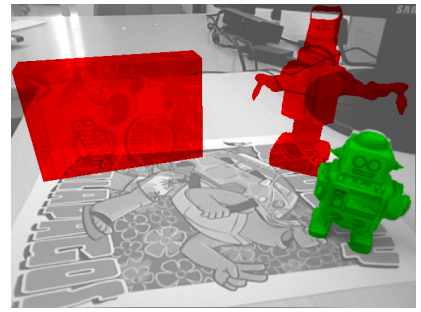
Fig. 8. Results from example objects: (a) Real images used for modeling the objects. (b)-(d) Final post-processed 3D meshes for these objects. A cereal box (see table under category "kitchen"), a full-sized PR2 (category "other"), a rubber duck, a small robot toy and a horse model (category toys).



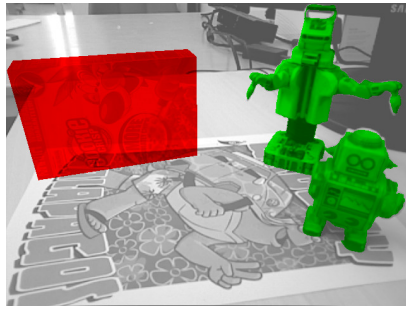
(a)



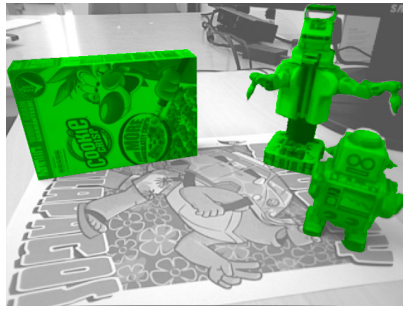
(b)



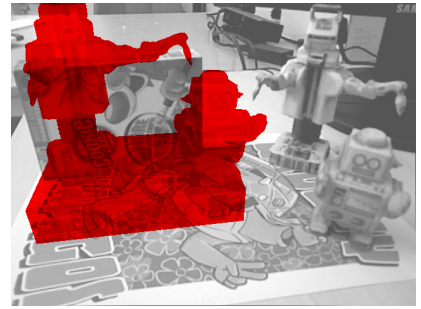
(c)



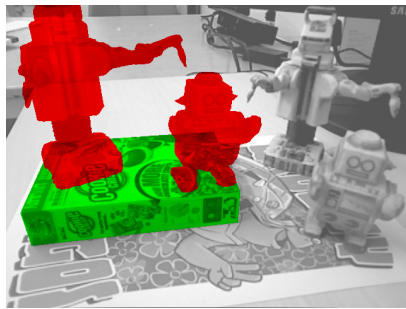
(d)



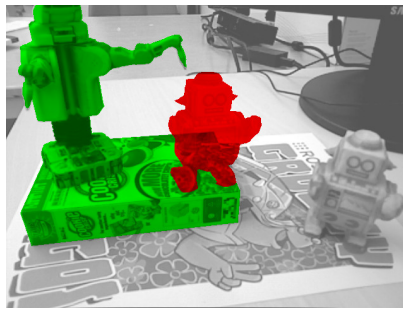
(e)



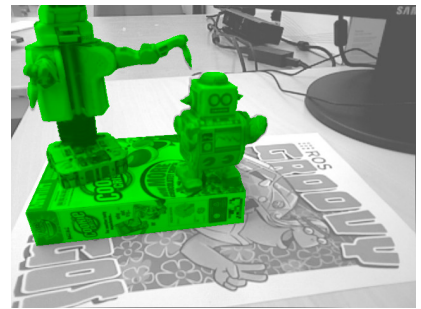
(f)



(g)



(h)



(i)

Fig. 9. Benchmarking using pre-defined scene layouts: (a) A marker is introduced in the scene, detected, and highlighted in blue. This marker provides a reference frame for the scene. (b) The desired object placement according to a pre-defined initial scene layout is highlighted in red. (c-e) One-by-one, the objects are placed in the scene and the color changes to green if their placement is sufficiently accurate. (f) A pre-defined target scene layout is highlighted in red. (g-i) The task is executed and the objects are moved to their target pose.

ID	Object Name	Dimension	Weight	Material	Other
Office					
1	Mead Index Cards	7.6x2x12.7 cm	136 g	paper in plastic	Amazon
2	Highland 6539 Self Stick Notes	11.7x5.3x4 cm	167.3 g	paper in plastic	Amazon
3	Paper Mate 12 Count Pencils	4.8x1.8x19.3 cm	68 g	cardboard	Amazon
4	Elmer's Washable No-Run School Glue	6.4x14x3 cm	45.36 g	plastic	Amazon
5	Keyboard	2.7x47.3x18.5	638 g	hard plastic	
6	Scissors	20.5x0.8x7.5 cm	60 g	hard plastic, metal	
7	Stapler	5x2.7x13 cm	136 g	hard plastic	
8	Hole Puncher	5.5x9x13.9 cm	483 g	metal	
9	Tipp-Ex	7.1x2.6x2.6 cm	60 g	hard plastic	
Kitchen					
10	First Years Take And Toss Straw Cup	8x8 x15 cm	141.7 g	plastic, paper	Amazon
11	Genuine Joe Plastic Stor Sticks	15x11.7x10.2 cm	249.5 g	cardboard	Amazon
12	Dr. Brown's Bottle Brush	5.3x9.7x31 cm	39.7 g	paper bottom with plastic	Amazon
13	Oreo mega stuff	20.3x5.1x15.2 cm	377 g	plastic	Amazon
14	Cheez It Big	31x16.5x16.5 cm	385.6 g	cardboard	Amazon
15	Cookie Crisp Box	29x6.8x19.3 cm	70 g	cardboard	empty
16	Plate	1.7x20.3x20.3 cm	351 g	porcelain	
17	Bowl	5.3x11.8x11.8 cm	80 g	porcelain	
18	Knife	21.5x1.3x1.8cm	25 g	steel	added texture
19	Fork	20x1.7x2.5cm	28 g	steel	added texture
Tools					
20	Stanley Piece Precision Screw-driver Set	19.6x9.9x2.3 cm	99.2 g	hard plastic	Amazon
21	Hammer	32.7x3.4x13.5 cm	658 g	hard plastic, metal	
22	Pitcher	16.4x2.8x5.3 cm	146 g	hard plastic	
23	Saw	63.5x2.5x4.3 cm	425 g	hard plastic	
24	Screwdriver	20x2.7x2.7 cm	56 g	hard plastic	
Toys					
25	6 Colored Highlighters	1.8x11.9x13.2 cm	39.7 g	plastic	Amazon
26	Crayola 64 Ct Crayons	14.5x12.7 cm	357.2 g	cardboard	Amazon
27	KONG Squeakair Tennis Ball with Rope Dog Toy	52.1x6.4x6.4 cm	82.2 g	textile	Amazon
28	Squeakin' Eggs Plush Dog Toys	17.8x7.6x14 cm	8.5 g	textile	Amazon
29	KONG Squeakair Sitting Duck Dog Toy	12.7x5.1x8.9 cm	8.5 g	textile, cardboard	Amazon
30	KONG Squeakair Sitting Frog Dog Toy	14x4.6x8.9 cm	8.5 g	textile, cardboard	Amazon
31	Munchkin White Hot Duck Bath Toy	13.2x7.1x9.7 cm	8.5 g	textile, cardboard	Amazon
32	Adventures of Huckleberry Finn	2x13x21.6 cm	181.44 g	paper	Amazon
33	Laugh-Out-Loud Jokes for Kids	10.7x0.8x17.3 cm	8.5 g	paper	Amazon
34	Black-White Duck	11.2x7.5x11.2 cm	110 g	rubber	
35	Small Robot	8.1x5.6x4.9 cm	60 g	metal	
36	Horse	10.3x3.2x15.7 cm	104 g	hard plastic	
37	Car	4.2x5.9x12.7 cm	145 g	metal	
Other					
38	PR2 Robot	178.9x118.6x136.9 cm	220 kg		
39	Mommys Helper Outlet Plugs	3.8x3.8x1.3 cm	22.7 g	plastic, cardboard	Amazon
40	Expo Dry Erase Board Eraser	5.3x13.2x3 cm	8.5 g	cardboard	Amazon

TABLE I
SUMMARY OF OBJECTS IN OUR INITIAL DATABASE RELEASE.

- [10] A. Kasper, Z. Xue, and R. Dillmann. "The KIT object models database: An object model database for object recognition, localization and manipulation in service robotics". In: *IJRR* 31.8 (2012), pp. 927–934.
- [11] B. Kehoe et al. "Cloud-based robot grasping with the google object recognition engine". In: *IEEE ICRA*. 2013, pp. 4263–4270.
- [12] D. Kragic, A. Miller, and P. Allen. "Real-time tracking meets online grasp planning". In: *IEEE ICRA*. Vol. 3. 2001, 2460–2465 vol.3.
- [13] Meshlab. <http://meshlab.sourceforge.net>. 2015.

- [14] A. T. Miller et al. "Automatic Grasp Planning Using Shape Primitives". In: *IEEE ICRA*. 2003, pp. 1824–1829.
- [15] K. Pauwels, L. Rubio, and E. Ros. "Real-time Pose Detection and Tracking of Hundreds of Objects". In: *IEEE Trans. on Circuits and Syst. for Video Tech.* (2015). , to appear.
- [16] K. Pauwels et al. "Real-time model-based rigid object pose estimation and tracking combining dense and sparse visual cues". In: *IEEE Comp. Soc. Conf. on Comp. Vision and Pattern Recog. (CVPR)*. Portland, Jan. 1, 2013, pp. 2347–2354.
- [17] F. T. Pokorny, Y. Bekiroglu, and D. Kragic. "Spherical Harmonics and Grasp Moduli Spaces". In: *IEEE ICRA*. 2014.
- [18] F. T. Pokorny, K. Hang, and D. Kragic. "Grasp Moduli Spaces". In: *Robotics: Science and Systems*. Berlin, Germany, 2013.
- [19] C. C. Print. <http://cubify.com/print/index>. 2015.
- [20] M. Przybylski et al. "A skeleton-based approach to grasp known objects with a humanoid robot". In: *IEEE-RAS Int. Conf. on Humanoid Robots (Humanoids)*. 2012, pp. 376–383.
- [21] J.-P. Saut and D. Sidobre. "Efficient models for grasp planning with a multi-fingered hand". In: *Robotics and Auton. Systems* (2012).
- [22] Sculpteo. <http://www.sculpteo.com/en/>. 2015.
- [23] Shapeways. <http://www.shapeways.com>. 2015.